BSDEs with Markov Chains and The Application: Homogenization of Systems of PDEs

Zhen WU

School of Mathematics,
Shandong University, Jinan 250100, China.
Email: wuzhen@sdu.edu.cn

joint with Dr. Huaibin TANG
INRIA-IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France

Brest, March, 2010

Main results in this talk:

we consider one kind of backward stochastic differential equations (BSDEs in short) where the coefficient f is affected by a Markovian switching.

1. Theoretical result:

1) We obtain the existence and uniqueness results for the solution to this kind of BSDEs.

2) We get the weak convergence result of BSDEs with a singular-perturbed Markov chain which is involved in a large state space.

2. Application: homogenization of one system of PDE.

Contents

Background

1. BSDEs:

Bismut (1973): stemmed BSDEs from stochastic control problem.

Pardoux & Peng (1990, 1991): general BSDEs driven by Brownian motion and probabilistic representation of PDE.

Tang & Li (1994), Barles, Buchdahn & Pardoux (1997): BSDEs with respect to both a Brownian motion and a Poisson random measure.

Nualart & Schoutens (2001), Bahlali, Eddahbi & Essaky (2003): BSDEs driven by a Lévy process.

2. Singularly perturbed Markov chains:

Zhang and Yin (1998): consider the Markov chain involved in a large state space, they introduced a small parameter ($\varepsilon > 0$) to highlight the contrast between the fast and slow transitions among different Markovian states and lead the Markov chain to a singularly perturbed one with two time-scale: the actual time t and the stretched time $\frac{t}{\varepsilon}$. As $\varepsilon \longrightarrow 0$, the asymptotic probability distribution of such Markov chain can be studied.

Zhang and Yin (1999, 2004): Applications in optimal control problem and mathematical finance: finding the **near-optimal control** of random-switching LQ optimal control problem, and **nearly optimal asset allocation** in hybrid stock investment models.

Our question:

BSDEs with singular-perturbed Markov chains and their application.

BSDEs with Markov Chains

$(\Omega, \mathcal{F}, P)$, $T > 0$, $\{\mathcal{H}_t, 0 \le t \le T\}$.

$M^2_{\mathcal{H}_t}(0, T; R^n)$: $\varphi = \{\varphi_t; t \in [0, T]\}$ satisfying $E \int_0^T |\varphi_t|^2 dt < \infty$;

$S^2_{\mathcal{H}_t}(0, T; R^n)$: $\varphi = \{\varphi_t; t \in [0, T]\}$ satisfying $E(\sup_{0 \le t \le T} |\varphi_t|^2) < \infty$.

B: $B_0 = 0$, d-dimensional $\mathcal{H}_t$-Brownian motion.

α: continuous-time Markov chain independent of B with the state space $\mathcal{M} = \{1, 2, \ldots, m\}$.

Let $\mathcal{N}$ denote the class of all P-null sets of $\mathcal{F}$. Denote $\mathcal{F}_t = \mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha \vee \mathcal{N}$.

Denote $M^2(0, T; R^n) = M^2_{\mathcal{F}_t}(0, T; R^n)$ and $S^2(0, T; R^n) = S^2_{\mathcal{F}_t}(0, T; R^n)$.

Remark: $\{\mathcal{F}_t; 0 \le t \le T\}$ is neither increasing nor decreasing, and it **does not constitute a filtration**.

Suppose the generator of the Markov chain $Q = (q_{ij})_{m\times m}$ is given by

$$P\{\alpha(t+\triangle) = j|\alpha(t) = i\} = \begin{cases} q_{ij}\triangle + o(\triangle), & \text{if } i \neq j, \\ 1 + q_{ij}\triangle + o(\triangle), & \text{if } i = j, \end{cases}$$

where $\triangle > 0$. Here $q_{ij} \geq 0$ is the transition rate from i to j if $i \neq j$, while $q_{ii} = -\sum_{j=1, i\neq j}^{m} q_{ij}$.

The generator Q is called **weakly irreducible** if the system of equations $\nu Q = 0$ and $\sum_{i=1}^{m} \nu_i = 1$ has a unique nonnegative solution. This nonnegative solution $\nu = (\nu_1, \cdots, \nu_m)$ is called the **quasi-stationary distribution** of Q.

Contents

Motivation

Some references can be seen **Zhang and Yin(1999)**, **Li and Zhou(2002)** and **Zhou and Yin(2003)**

Consider the following stochastic LQ control problem with Markov jumps

$$\text{min } J(v(\cdot)) = \frac{1}{2}E\left\{\int_0^T [x_t' R(t,\alpha_t)x_t + v_t' N(t,\alpha_t)v_t]dt + x_T' Q(\alpha_T)x_T\right\}$$

$$\text{s.t. } \begin{cases} dx_t = [A(t,\alpha_t)x_t + B(t,\alpha_t)v_t]\,dt + [C(t,\alpha_t)x_t + D(t,\alpha_t)v_t]dB_t, \\ x_0 = x \in R^n, \end{cases}$$

Admissible controls set: $\mathcal{U}_{ad} \equiv M^2_{\mathcal{H}_t}(0,T;R^{n_u \times d})$.

Our aim is to find an admissible control $u(\cdot)$ such that

$$J(u(\cdot)) = \inf_{v \in \mathcal{U}_{ad}} J(v(\cdot)).$$

Here, we use FBSDE to derive its optimal control:

Theorem 2.1 For any admissible control $v(\cdot)$, if the following FBSDE admits a unique solution (x_t^v, y_t^v, z_t^v)

$$\begin{cases} dx_t^v = [A(t,\alpha_t)x_t^v + B(t,\alpha_t)v_t]dt + [C(t,\alpha_t)x_t^v + D(t,\alpha_t)v_t]dB_t, \\ -dy_t^v = [A'(t,\alpha_t)y_t^v + C'(t,\alpha_t)z_t^v + R(t,\alpha_t)x_t^v]dt - z_t^v dB_t, \\ x_0^v = x, \quad y_T^v = Q(\alpha_T)x_T^v \end{cases}$$

Then there exists a unique optimal control for the above LQ problem

$$u(t) = -N^{-1}(t,\alpha_t)[B'((t,\alpha_t))y_t + D'(t,\alpha_t)z_t].$$

Here (x_t, y_t, z_t) is the solution of FBSDE with respect to the control $u(\cdot)$.

Contents

BSDEs with Markov Chains

Firstly, we will study the following BSDEs with Markov chains

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, \alpha_s)ds - \int_t^T Z_s dB_s. \qquad (1)$$

Assumption 2.1

(i) $\xi \in L^2(\mathcal{F}_T; R^k)$;

(ii) $f : \Omega \times [0,T] \times R^k \times R^{k\times d} \times \mathcal{M} \to R^k$ satisfies that $\forall\ (y,z) \in R^k \times R^{k\times d}$, $\forall i \in \mathcal{M}$, $f(\cdot, y, z, i) \in M^2_{\mathcal{F}_t^B}(0,T; R^k)$, and $f(t,y,z,i)$ is uniformly Lipschitz continuous with respect to y and z, i.e., $\exists\ \mu > 0$, such that $\forall\ (\omega, t) \in \Omega \times [0,T]$, (y_1, z_1), $(y_2, z_2) \in R^k \times R^{k\times d}$,

$$|f(t, y_1, z_1, i) - f(t, y_2, z_2, i)| \le \mu(|y_1 - y_2| + |z_1 - z_2|).$$

Theorem 2.2 Under Assumption 2.1, there is a unique solution pair $(Y, Z) \in S^2(0, T; R^k) \times M^2(0, T; R^{k\times d})$ for BSDE (1).

The following **extension of Itô's martingale representation theorem** and its corollary play key role during the proof of this theorem.

Proposition 2.1 Define a filtration $(\mathcal{G}_t)_{0\le t\le T}$ by $\mathcal{G}_t = \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha$ where α and B are independent with each other. For $M \in L^2(\mathcal{G}_T; R^k)$, there exists a unique random variable $M_0 \in L^2(\mathcal{F}_T^\alpha; R^k)$ and a unique stochastic process $Z = \{Z_t; 0 \le t \le T\} \in M^2_{\mathcal{G}_t}(0, T; R^{k\times d})$ such that

$$M = M_0 + \int_0^T Z_t dB_t, \quad 0 \le t \le T.$$

Actually, $M_0 = E[M|\mathcal{F}_T^\alpha]$.

Corollary 2.1 For $t \leq T$, we consider the filtration $(\mathcal{N}_s)_{t\leq s\leq T}$ defined by $\mathcal{N}_s = \mathcal{F}_s^B \vee \mathcal{F}_{t,T}^{\alpha}$. For $M \in L^2(\mathcal{N}_T; R^k)$, there exist $M_t \in L^2(\mathcal{F}_{t,T}^{\alpha} \vee \mathcal{F}_t^B; R^k)$ and a unique stochastic process $Z = \{Z_s; t \leq s \leq T\} \in M^2_{\mathcal{N}_s}(t,T; R^{k\times d})$ such that

$$M = M_t + \int_t^T Z_s dB_s.$$

Proof of Theorem 2.2 Firstly, we will introduce a new filtration. Define the filtration $(\mathcal{G}_t)_{0\leq t\leq T}$ by

$$\mathcal{G}_t \triangleq \mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha} \vee \mathcal{N}.$$

Combing with extension of Itô's martingale representation theorem (**Proposition 2.1**), we can show that BSDE

$$Y_t = \xi + \int_t^T f(s, \alpha_s) ds - \int_t^T Z_s dB_s, \qquad 0 \le t \le T,$$

has a solution pair $(Y, Z) \in S^2(0, T; R^k) \times M^2(0, T; R^{k \times d})$.

One technical difficulty is to prove that the processes $Y = \{Y_t; 0 \le t \le T\}$ and $Z = \{Z_t; 0 \le t \le T\}$ are $\mathcal{F}_t$-measurable, i.e. $\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$-measurable. We used Doob's martingale convergence theorem and gave the careful analysis for the σ-algebra generated by the (discrete time sequences) Markov chains and Brownian motion, then get the desired conclusion. Similar technique can be seen in **Pardoux and Peng(1994)**.

With the fixed point method, we can get the corresponding result for the general case. □

As a corollary, we give an estimation of the solution.

Corollary 2.2 Under Assumption 2.1, we have the following estimation for the solution of BSDE (1)

$$E\left(\sup_{0\le t\le T}|Y_t|^2+\int_0^T Z_t^2dt\right)\le CE\left(|\xi|^2+\int_0^T|f(t,0,0,\alpha_t)|^2dt\right).$$

Applying Itô's formula, Schwartz's inequality, Gronwall's lemma and Burkholder-Davis-Gundy inequality, we can get the proof.

Contents

Preliminary of Singularly Perturbed Markov Chains

Consider the case that the Markov chain has a large state space which can be divided into a number of weakly irreducible classes such that it fluctuates rapidly among different states in a weakly irreducible class, and jumps less frequently among weakly irreducible classes.

To distinguish the fast transitions from the slow transitions among different states, **Zhang & Yin (1998)** showed that one can introduce a small parameter $\varepsilon > 0$ which leads to a singularly perturbed system involving two-time scales (actual time t and the stretched time $\frac{t}{\varepsilon}$).

For a continuous-time ε-dependent singularly perturbed Markov chain $\alpha^\varepsilon = \{\alpha_t^\varepsilon; 0 \le t \le T\}$ which have the generator $Q^\varepsilon = \frac{1}{\varepsilon}\tilde{Q} + \hat{Q}$, where $\tilde{Q}$ and $\hat{Q}$ are time-invariant generators. Here $\tilde{Q} = \mathrm{diag}(\tilde{Q}^1, \cdots, \tilde{Q}^l)$. For $k \in \{1, \cdots, l\}$, $\tilde{Q}^k$ is the weakly irreducible generator corresponding to the states in $\mathcal{M}_k = \{s_{k1}, \cdots, s_{km_k}\}$. The state space can be decomposed as $\mathcal{M} = \{1, 2, \cdots, m\} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l$.

The generator $\tilde{Q}$ dictates the fast motion of the Markov chain and $\hat{Q}$ governs the slow motion. That means that the slow and fast components are coupled through weak and strong interactions in the sense that the underlying Markov chain fluctuates rapidly in a single group $\mathcal{M}_k$ and jumps less frequently among groups $\mathcal{M}_k$ and $\mathcal{M}_j$ for $k \neq j$. The states in $\mathcal{M}_k$, $k = 1, \cdots, l$, are not isolated or independent of each other.

More precisely, if we consider the states in $\mathcal{M}_k$ as a "single" state, then the transition rate between these "states" are described by the element of matrix $\hat{Q}$.

As one aggregate the states in $\mathcal{M}_k$ as a single state, all such states are coupled by $\hat{Q}$. So we can define an **aggregated process** $\bar{\alpha}^{\varepsilon} = \{\bar{\alpha}_t^{\varepsilon}; 0 \leq t \leq T\}$ by $\bar{\alpha}_t^{\varepsilon} = k$, when $\alpha_t^{\varepsilon} \in \mathcal{M}_k$.

As shown in the following **Proposition 3.1**, the process $\bar{\alpha}^{\varepsilon}$ is not necessarily Markovian, but it converges weakly to a continuous time Markov chain $\bar{\alpha}$.

Proposition 3.1 (Zhang and Yin(1998))
(i) $\bar{\alpha}^\varepsilon$ converges weakly to $\bar{\alpha}$ generated by

$$\bar{Q} = \text{diag}(\nu^1, \cdots, \nu^l)\hat{Q}\text{diag}(\mathbb{I}_{m_1}, \cdots, \mathbb{I}_{m_l}),$$

as $\varepsilon \to 0$, where ν^k is the quasi-stationary distribution of $\tilde{Q}^k$, $k = 1, \cdots, l$, and $\mathbb{I}_k = (1, \cdots, 1)' \in R^k$.
(ii) For any bounded deterministic function $\beta(\cdot)$,

$$E\left[\int_s^T (I_{\{\alpha_t^\varepsilon = s_{kj}\}} - \nu_j^k I_{\{\bar{\alpha}_t^\varepsilon = k\}})\beta(t)dt\right]^2 = O(\varepsilon),$$

$\forall\ k = 1, \cdots, l, \forall\ j = 1, \cdots, m_k$.
Here I_A is the indicator function of a set A.

We can see that $\hat{Q}$ together with the quasi-stationary distributions of $\tilde{Q}^k$, $k = 1, \cdots, l$, determine the transition's probability among states in $\mathcal{M}_k$, for $k = 1, \cdots, l$. The probability distribution of the underlying Markov chain will quickly reach a stationary regime determined by $\tilde{Q}$. And the influence of $\hat{Q}$ takes over subsequently.

As an obviously result of Lemma 7.3 in **Zhang & Yin (2004)**, we have

Proposition 3.2 Suppose (i) $g(t,x)$ is a function defined on $[0,T]\times R^m$ satisfying that $g(\cdot,\cdot)$ is Lipschitz continuous with x and $\forall\ x\in R^m$, either $|g(t,x)|\leq K(1+|x|)$ or $|g(t,x)|\leq K$; (ii) a sequence of stochastic process indexed by ε, $\{x_t^\varepsilon; 0\leq t\leq T\}$ is in $M^2_{\mathcal{F}_t}(0,T;R^m)$ and there exists a constant $C>0$ such that $E(\sup_{0\leq t\leq T}|x_t^\varepsilon|^2)\leq C$.

Denote $\pi_{ij}^\varepsilon(t)=\pi_{ij}^\varepsilon(t,\alpha_t^\varepsilon)$, with $\pi_{ij}^\varepsilon(t,\alpha)=I_{\{\alpha=s_{ij}\}}-\nu_j^i I_{\{\alpha\in M_i\}}$, then for any $k=1,\cdots,l, j=1,\cdots,m_k$,

$$\sup_{0<t\leq T} E\left|\int_0^t g(t,x_s^\varepsilon)\pi_{ij}^\varepsilon(s,\alpha_s^\varepsilon)ds\right|\rightarrow 0,\quad \text{as } \varepsilon\rightarrow 0.$$

Contents

BSDEs with singularly perturbed Markov chains

Denote $D(0,T;R^k)$ be the space of càdlàg trajectories endowed with the " Meyer-Zheng" topology, i.e., the topology of convergence in dt-measure.

Consider the following BSDE with a singularly perturbed Markov chain

$$Y_t^\varepsilon = \xi + \int_t^T f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)ds - \int_t^T Z_s^\varepsilon dB_s,$$

Set $M_t^\varepsilon = \int_0^t Z_s^\varepsilon dB_s$, we can rewrite the above BSDE as

$$Y_t^\varepsilon = \xi + \int_t^T f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)ds - (M_T^\varepsilon - M_t^\varepsilon), \tag{2}$$

Our aim: show that $(Y^\varepsilon, M^\varepsilon)$ converges weakly "in the sense of Meyer-Zheng" if we equip the space of paths with the topology of convergence in dt-measure.

Since it is hard to show that the sequence (Z^ε) is tight, here, we will restrict that the generator f does not depend on Z.

Firstly, we make the following assumption:

Assumption 3.1
(i) $\xi \in L^2(\mathcal{F}_T^B; R^k)$.
(ii) For $f : [0,T] \times R^k \times \mathcal{M} \to R^k$, there exists a constant $C > 0$ such that $\sup\limits_{\substack{0 \le t \le T \\ 1 \le i \le m}} |f(t,0,i)| \le C$.

Our main result here is the following theorem:

Theorem 3.1 Under Assumption 2.1 and Assumption 3.1, the sequence of process $(Y_t^\varepsilon, \int_t^T Z_s^\varepsilon dB_s)$ converges in law to the process $(Y_t, \int_t^T Z_s dB_s)$ as $\varepsilon \to 0$, when probability measures on $C(0,T;R^{2k})$ equipped with the topology of convergence in dt measure. Here (Y,Z) is the solution pair to the following BSDE with an averaged Markov chain

$$Y_t = \xi + \int_t^T \bar{f}(s, Y_s, \bar{\alpha}_s)ds - \int_t^T Z_s dB_s, \tag{3}$$

$\bar{\alpha}$ is the averaged Markov chain, and $\bar{f}(s,y,i) = \sum_{j=1}^{m_i} \nu_j^i f(t,y,s_{ij})$ for $i \in \bar{\mathcal{M}} = \{1, \cdots, l\}$. Moreover, as $\varepsilon \to 0$, the $\mathcal{F}_T^{\alpha^\varepsilon}$-measurable random variables sequence (Y_0^ε) converges in law to the random variable Y_0 which is $\mathcal{F}_T^{\bar{\alpha}}$-measurable.

This proof is consisted of two steps.

Step 1: Tightness and convergence for $(Y^\varepsilon, M^\varepsilon)$.

Proposition 3.3 Under Assumption 2.1 and Assumption 3.1, BSDEs (2) and (3) have unique solutions $(Y^\varepsilon, Z^\varepsilon)$ and $(Y, Z) \in S^2(0,T;R^k) \times M^2(0,T;R^{k\times d})$, and there exists a positive constant C which does not depend on ε, such that

$$E\left[\sup_{0\le t\le T}|Y_t^\varepsilon|^2 + \int_0^T (Z_t^\varepsilon)^2 dt\right] \le C,$$

$$E\left[\sup_{0\le t\le T}|Y_t|^2 + \int_0^T (Z_t)^2 dt\right] \le C.$$

We need the following "Meyer-Zheng tightness criteria" given in **Meyer & Zheng (1984)**:

Lemma 3.1 The sequence of semi-martingale $\{V_s^n; 0 \le s \le t\}$ defined on the filtered probability spaces $(\Omega, \{\mathcal{F}_s, 0 \le s \le t\}, \mathcal{F}, P^n)$ is tight whenever

$$\sup_n \left(\sup_{0 \le s \le t} E^n |V_s^n| + CV_t(V^n) \right) < \infty,$$

where $CV_t(V^n)$ denotes the "conditional variation of V^n on $[0, t]$" defined by

$$CV_t(V^n) = \sup E^n \left(\sum_i |E^n(V_{t_{i+1}}^n - V_{t_i}^n | \mathcal{F}_{t_i}^n)| \right),$$

with "sup" meaning the supremum has taken over all partitions of the interval $[0, t]$.

Proposition 3.4 The sequence of $(Y^\varepsilon, M^\varepsilon)$ is tight on the space $D(0,T;\, R^k) \times D(0,T;R^k)$.

Proof: Let $\mathcal{G}_t^\varepsilon = \mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha^\varepsilon} \vee \mathcal{N}$, we define

$$CV_t(Y^\varepsilon) = \sup E\left(\sum_i |E(Y^\varepsilon_{t_{i+1}} - Y^\varepsilon_{t_i}|\mathcal{G}^\varepsilon_{t_i})|\right),$$

where the supreme is over all partitions of the interval $[0,T]$.

Since M^ε is a $\mathcal{G}_t^\varepsilon$-martingale, it follows that

$$CV_t(Y^\varepsilon) \leq E\int_0^T |f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)|ds.$$

Thus, there exists a constant $C > 0$ which is independent of ε, such that

$$\sup_{\varepsilon} E[\sup_{0\le t\le T} |Y_t^\varepsilon|^2 + < M_t^\varepsilon >] \le C.$$

It follows that

$$\sup_{\varepsilon}\left(CV_t(Y^\varepsilon) + \sup_{0\le t\le T} |Y_t^\varepsilon| + \sup_{0\le t\le T} |M_t^\varepsilon| \right) < \infty.$$

Hence the sequences (Y^ε) and (M^ε) satisfy the "Meyer-Zheng tightness criteria" and the result is followed.□

Step 2: Identification of the limit.

From **Proposition 3.4**, we know that there exists a subsequence of $(Y^\varepsilon, M^\varepsilon)$, which still denoted by $(Y^\varepsilon, M^\varepsilon)$, and which converges in law on the space $D(0,T;R^k) \times D(0,T;R^k)$ towards a càdlàg process (Y, M). Furthermore, there exists a countable subset D of $[0,T]$, such that $(Y^\varepsilon, M^\varepsilon)$ converges in finite-distribution to (Y, M) on D^c.

Proposition 3.5 For the limit process (Y, M), we have
(i) For every $t \in [0, T] - D$,

$$Y_t = \xi + \int_t^T \bar{f}(s, Y_s, \bar{\alpha}_s)ds - (M_T - M_t).$$

(ii) M is a $\mathcal{H}_t$-martingale, where $\mathcal{H}_t = \mathcal{F}_t^B \vee \mathcal{F}_T^{\bar{\alpha}}$.
Proof: as $\varepsilon \to 0$, from **Proposition 3.2**,

$$\sup_{0 \le t \le T} \left| E \int_0^t f(s, Y_s^\varepsilon, s_{ij})[I_{\{\alpha_s^\varepsilon = s_{ij}\}} - \nu_j^i I_{\{\alpha_s^\varepsilon \in \mathcal{M}_i\}}]ds \right| \to 0, .$$

Since as $\varepsilon \to 0$, on $C(0, T; R^k)$

$$\int_0^\cdot \bar{f}(s, Y_s^\varepsilon, \bar{\alpha}_s)ds \text{ converges weakly to } \int_0^\cdot \bar{f}(s, Y_s, \bar{\alpha}_s)ds,$$

and

$$
\begin{aligned}
&\int_0^t f(s, Y_s^\varepsilon, \alpha_s^\varepsilon) ds \\
&= \int_0^t \sum_{i=1}^{l} \sum_{j=1}^{m_i} f(s, Y_s^\varepsilon, s_{ij}) I_{\{\alpha_s^\varepsilon = s_{ij}\}} \\
&= \int_0^t \sum_{i=1}^{l} \sum_{j=1}^{m_i} f(s, Y_s^\varepsilon, s_{ij}) [I_{\{\alpha_s^\varepsilon = s_{ij}\}} - \nu_j^i I_{\{\alpha_s^\varepsilon \in \mathcal{M}_i\}}] ds \\
&\quad + \int_0^t \bar{f}(s, Y_s^\varepsilon, \bar{\alpha}_s) ds,
\end{aligned}
$$

passing to the limit in the BSDE (2) and we can derive assertion (i).

Now, we prove assertion (ii).

For any $0 \leq t_1 \leq t_2 \leq T$, Φ_{t_1} is a continuous mapping from $C(0,t_1;R^d) \times D(0,t_1;R^k) \times D(0,T;\bar{\mathcal{M}})$. $\forall\ \varepsilon > 0$, since M^ε is a martingale with respect to $\mathcal{G}_t^\varepsilon = \mathcal{F}_T^{\alpha^\varepsilon} \vee \mathcal{F}_t^B$, Y^ε and $\bar{\alpha}^\varepsilon$ are $\mathcal{G}_t^\varepsilon$-adapted, we know

$$E\left[\Phi_{t_1}(B, Y^\varepsilon, \bar{\alpha}^\varepsilon) \int_0^\delta (M_{t_2+r}^\varepsilon - M_{t_1+r}^\varepsilon) dr\right] = 0,$$

here B is the Brownian motion. From the weak convergence of $(Y^\varepsilon, \bar{\alpha}^\varepsilon)$ and the fact that $E(\sup_{0\leq t\leq T} |M_t^\varepsilon|^2) \leq C$, we obtain

$$E\left[\Phi_{t_1}(B, Y, \bar{\alpha}) \int_0^\delta (M_{t_2+r} - M_{t_1+r}) dr\right] = 0.$$

Dividing the second identity by δ, letting $\delta \to 0$, and exploiting the right continuity, we obtain that

$$E\left[\Phi_{t_1}(B, Y, \bar{\alpha})(M_{t_2} - M_{t_1})\right] = 0.$$

From the freedom choice of t_1, t_2 and Φ_{t_1}, we deduce that M is a $\mathcal{H}_t$-martingale. $\square$

We also can get

Proposition 3.6 Let $\{(\bar{Y}_t, \bar{Z}_t); 0 \le t \le T\}$ be the unique solution of BSDE (3), then $\forall\ t \in [0, T]$,

$$E|Y_t - \bar{Y}_t|^2 + E\left([M - \int_0^\cdot \bar{Z}_r dB_r]_T - [M - \int_0^\cdot \bar{Z}_r dB_r]_t\right) = 0.$$

Since $\bar{Y}$ is continuous, Y is c$\grave{a}$dl$\grave{a}$g and D is countable, we get $Y_t = \bar{Y}_t$, $P-$a.s., $\forall\ t \in [0, T]$. Moreover, we can deduce that $M \equiv \bar{M}$. Hence, we get the result that the sequence $(Y_t^\varepsilon, \int_t^T Z_s^\varepsilon dB_s)$ converges in law to the process $(Y_t, \int_t^T Z_s dB_s)$, and the proof of **Theorem 3.1** is completed.□

Example 3.1: Consider the case that $\tilde{Q}$ is weakly irreducible with the state space $\mathcal{M} = \{1, \cdots, m\}$ and $\nu = (\nu_1, \cdots, \nu_m)$ is the quasi-stationary distribution, the corresponding BSDE is

$$Y_t^\varepsilon = \xi + \int_t^T f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)ds - \int_t^T Z_s^\varepsilon dB_s. \tag{4}$$

From our results, as $\varepsilon \to 0$, the sequence of process $(Y_t^\varepsilon, \int_t^T Z_s^\varepsilon dB_s)$ converges in law to the process $(Y_t, \int_t^T Z_s dB_s)$, where (Y, Z) is the unique solution to the following BSDE

$$Y_t = \xi + \int_t^T \sum_{i=1}^m \nu_i f(s, Y_s, i)ds - \int_t^T Z_s dB_s. \tag{5}$$

It is noted that the generator of BSDE (5) depends on the quasi-stationary distribution of the Markov chain, instead of depending on the Markov chain. i.e., Y, which is the asymptotic solution of $\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^{\alpha^\varepsilon}$-adapted process Y^ε, is $\mathcal{F}_t^B$-adapted.

Example 3.2: Suppose the generator of the continuous-time Markov chain affected BSDE (1) is $Q = \begin{pmatrix} -22 & 20 & 2 \\ 41 & -42 & 1 \\ 1 & 2 & -3 \end{pmatrix}$, and the corresponding state space is $\mathcal{M} = \{s_1, s_2, s_3\}$. It is obvious that the transition rate between s_1 and s_2 is larger than the transition rate between s_3 and other states, i.e., the jumps between s_1 and s_2 are more frequent than jumps between s_3 and other states. We can rewrite Q as following:

$$Q = \frac{1}{0.05}\tilde{Q} + \hat{Q} = \frac{1}{0.05}\begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{pmatrix}$$

It is noted that we choose suitable ε to guarantee that $\tilde{Q}$ and $\hat{Q}$ to be the generator with the same order of magnitude. Introduce the continuous-time ε-dependent singularly perturbed Markov chain $\alpha^\varepsilon = \{\alpha_t^\varepsilon; 0 \le t \le T\}$ which have the generator

$$Q^\varepsilon = \frac{1}{\varepsilon}\tilde{Q} + \hat{Q} = \frac{1}{\varepsilon}\begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{pmatrix},$$ and

define the aggregated process

$$\bar{\alpha}^\varepsilon = \{\bar{\alpha}_t^\varepsilon; 0 \le t \le T\} = \begin{cases} 1, & \alpha_t^\varepsilon \in \{s_1, s_2\}, \\ 2, & \alpha_t^\varepsilon \in \{s_3\}. \end{cases}$$

Result of **Zhang & Yin (1998)** (**Proposition 3.1**) yields that $\bar{\alpha}^\varepsilon$ converges weakly to a continuous-time Markov chain $\bar{\alpha}$ generated by $\bar{Q} = \begin{pmatrix} -\frac{5}{3} & \frac{5}{3} \\ 3 & -3 \end{pmatrix}$. From our result, we can adopt the solution of the following BSDE

$$Y_t = \xi + \int_t^T \bar{f}(s, Y_s, \bar{\alpha}_s)ds - \int_t^T Z_s dB_s,$$

as an asymptotic solution of the original BSDE.

Remark: It is noted that in practical systems, the small parameter ε is just a fixed parameter and it separates different scales in the sense of order of magnitude in the generator. It does not need to tend to 0.

Background

Pardoux & Peng (1992): BSDEs provide a probabilistic representation for the solution of PDE.

Then BSDEs provide a probabilistic tool to study the **homogenization of PDEs**, which is the process of replacing rapidly varying coefficients by new ones thus the solutions are close. It is noted that there are two different probabilistic schemes based on BSDEs.

Briand & Hu (1998), Buckdahn, Hu & Peng (1999): based on a stability property of BSDEs.

Pardoux (1999), Essaky & Ouknine (2006), Bahlali, Elouaflin & Pardoux (2009): based on the theory of weak convergence of BSDEs under the "topology of Meyer-Zheng" (much weaker than Skorohod's topology), i.e., the topology of convergence in dt-measure.

Main results

For $x \in R^m$, consider the following sequence of semi-linear backward PDE with a singularly perturbed Markov chain, indexed by $\varepsilon > 0$,

$$u^\varepsilon(t,x) = h(x) + \int_t^T [\mathcal{L}u^\varepsilon(r,x) + f(r,x,u^\varepsilon(r,x),\alpha_r^\varepsilon)]dr, \quad 0 \le t \le T.$$

Here α^ε is a singularly perturbed Markov chain,
$\mathcal{L}u = (Lu_1, \cdots, Lu_k)'$, with $L = \frac{1}{2}\sum_{i,j=1}^{m} (\sigma\sigma')_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(t,x)\frac{\partial}{\partial x_i}$.

We have the following homogenization result:

Theorem 4.1 Under suitable assumptions , the above PDE has a classical solution $\{u^\varepsilon(t,x); 0 \le t \le T, x \in R^m\}$. As $\varepsilon \to 0$, the sequence of u^ε converges in law to a process u, where $u(t,x)$ is the classical solution of the following PDE with an averaged Markov chain

$$u(t,x) = h(x) + \int_t^T [\mathcal{L}u(r,x) + \bar{f}(r,x,u(r,x),\bar{\alpha}_r)]dr, \quad 0 \le t \le T.$$

Here $\bar{\alpha}$ is the averaged Markov chain and $\bar{f}$ is the average of f defined as $\bar{f}(t,x,u,i) = \sum_{j=1}^{m_i} \nu_j^i f(t,x,u,s_{ij})$, for $i \in \bar{\mathcal{M}} = \{1,\cdots,l\}$.

Relation between semi-linear PDEs and BSDEs with Markov chains

The main part of the proof is to prove the relation between semi-linear PDEs and BSDEs with Markov chains.

$C^k(R^p; R^q)$: space of functions of class C^k from R^p to R^q,

$C^k_{l,b}(R^p; R^q)$: space of functions of class C^k whose partial derivatives of order less than or equal to k are bounded,

$C^k_p(R^p; R^q)$: space of functions of class C^k which, together with all their partial derivatives of order less than or equal to k, grow at most like a polynomial function of the variable x at infinity.

Consider the following semi-linear backward PDE on $0 \leq t \leq T$,

$$u(t,x) = h(x) + \int_t^T [\mathcal{L}u(r,x) + f(r,x,u(r,x),(\nabla u\sigma)(r,x),\alpha_r)]dr, \quad (6)$$

Firstly, we make the following assumption:

Assumption 4.1 $b \in C^3_{l,b}(R^m; R^m)$, $\sigma \in C^3_{l,b}(R^m; R^{m\times d})$, $h \in C^3_p(R^m; R^k)$. $f : [0,T] \times R^m \times R^k \times R^{k\times d} \times \mathcal{M} \to R^k$, $\forall\ s \in [0,T]$, $\forall\ i \in \mathcal{M}$, $(x,y,z) \to f(s,x,y,z,i)$ is of class C^3.

Moreover, $f(s,\cdot,0,0,i) \in C^3_p(R^m; R^k)$, and its first and second order partial derivatives in y and z are bounded on $[0,T] \times R^m \times R^k \times R^{k\times d} \times \mathcal{M}$, as well as its derivatives of order one and two with respect to x.

Definition 4.1 A classical solution of PDE (6) is a R^k-valued stochastic process $\{u(t,x); 0 \le t \le T, x \in R^m\}$ which is in $C^{0,2}([0,T] \times R^m; R^k)$ and satisfies that $u(t,x)$ is $\mathcal{F}^{\alpha}_{t,T}$-measurable, for all (t,x).

$\forall\ t \in [0, T]$, $x \in R^m$, we introduce the following SDE and BSDE with a Markov chain on $t \leq s \leq T$

$$X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dB_r, \tag{7}$$

$$Y_s^{t,x} = h(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}, \alpha_r)dr - \int_s^T Z_r^{t,x} dB_r \tag{8}$$

For $s \leq t$, define $X_s^{t,x} = X_{s\vee t}^{t,x}$, $Y_s^{t,x} = Y_{s\vee t}^{t,x}$ and $Z_s^{t,x} = 0$, then $(X, Y, Z) = (X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})$ is defined on $(s, t) \in [0, T]^2$.

We can show that, under above assumptions, the FBSDE (7)-(8) provides both a probabilistic representation and the unique classical solution to PDE (6).

Theorem 4.2 Under Assumption 4.1, let $\{u(t,x); 0 \leq t \leq T, x \in R^m\}$ be a classical solution of PDE (6), and suppose that there exists a constant C such that,

$$|u(t,x)| + |\partial_x u(t,x)\sigma(t,x)| \leq C(1+|x|), \quad \forall\ (t,x) \in [0,T] \times R^m,$$

then $\left(Y_s^{t,x} = u(s, X_s^{t,x}), Z_s^{t,x} = \partial_x u(s, X_s^{t,x})\sigma(s, X_s^{t,x}); t \leq s \leq T\right)$ is the unique solution of BSDE (8), here $(X_s^{t,x}; t \leq s \leq T)$ is the solution to SDE (7).

Now we will deduce the converse side of Theorem 4.2.

Theorem 4.3 Suppose that for some $p > 0$, $E|h(x)|^p$ $+E\int_0^T |f(t,x,0,0,\alpha)|dt < \infty$, then under Assumption 4.1, the process $\{u(t,x) = Y_t^{t,x}; 0 \le t \le T, x \in R^m\}$ is the unique classical solution to PDE (6).

The proof is mainly depends on the following two propositions about the regularity of the solution of BSDE (8).

Proposition 4.1 $\{Y_s^{t,x}; (s,t) \in [0,T]^2, x \in R^m\}$ has a version whose trajectories belong to $C^{0,0,2}([0,T]^2 \times R^m)$. Hence for all $t \in [0,T]$, $x \to Y_t^{t,x}$ is of class C^2 a.s..

Proposition 4.2 $\{Z_s^{t,x}; (s,t) \in [0,T]^2, x \in R^m\}$ has an a.s. continuous version which is given by $Z_s^{t,x} = \nabla Y_s^{t,x}(\nabla X_s^{t,x})^{-1}\sigma(X_s^{t,x})$. In particular, $Z_t^{t,x} = \nabla Y_t^{t,x}\sigma(x)$. Here $\left(\nabla Y_s^{t,x} = \dfrac{\partial Y_s^{t,x}}{\partial x}, \nabla Z_s^{t,x} = \dfrac{\partial Z_s^{t,x}}{\partial x}\right)$ is the unique solution to

$$\begin{aligned}\nabla Y_s^{t,x} =& h'(X_T^{t,x})\nabla X_T^{t,x} + \int_s^T [f'_x(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}, \alpha_r)\nabla X_r^{t,x} \\ &+ f'_y(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}, \alpha_r)\nabla Y_r^{t,x} \\ &+ f'_z(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}, \alpha_r)\nabla Z_r^{t,x}]dr - \int_s^T Z_r^{t,x} dB_r.\end{aligned}$$

Thank you!